BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 6: e94416
doi: 10.3897/biss.6.94416

OPEN ACCESS

Conference Abstract

Community Data Mobilization in Wikidata: A paleontology perspective

Holly Little[‡], Talia Karim[§], Erica Krimmel[|], Ben Norton[¶], Jessica Utrup[#], Lindsay J. Walker[¤], Jacob Van Veldhuizen[§]

‡ Smithsonian National Museum of Natural History, Washington, DC, United States of America
§ University of Colorado, Boulder, Colorado, United States of America
| iDigBio, Florida State University, Tallahassee, Florida, United States of America
¶ North Carolina Museum of Natural Sciences, Raleigh, North Carolina, United States of America
# Yale Peabody Museum, New Haven, Connecticut, United States of America
¤ Symbiota, Arizona State University, Tempe, Arizona, United States of America

Corresponding author: Holly Little (littleh@si.edu)

Received: 02 Sep 2022 | Published: 07 Sep 2022

Citation: Little H, Karim T, Krimmel E, Norton B, Utrup J, Walker LJ, Van Veldhuizen J (2022) Community Data Mobilization in Wikidata: A paleontology perspective. Biodiversity Information Science and Standards 6: e94416. https://doi.org/10.3897/biss.6.94416

Abstract

Wikidata offers a centralized, accessible platform for working collaboratively to disambiguate entities, e.g., people associated with biodiversity collections, and to mobilize information about them. This shared information can be used to improve connections across institutions and individuals, to augment local data records, and to encourage expertise-sharing. Over the past year, the Paleo Data Working Group has organized several events to bring the Wikidata movement to paleontological collections. This talk will share key products and findings generated by these efforts, and note parallels to disciplines beyond paleontology.

In the global collections community, momentum has been building to leverage linked data principles and facilitate data discovery using the Wikidata platform. An example of this is the increasing adoption of Wikidata specifically for the purpose of storing biographical information about people associated with collections, because doing so facilitates discovery about who is doing what work across multiple institutions, and with what impact (Groom et al. 2020, Güntsch et al. 2021). In March 2022, thirty participants gathered virtually for a participatory workshop, *Using Wikidata to capture and share information*

This is an open access article distributed under the terms of the CC0 Public Domain Dedication.

about people in paleontology. This workshop was an introduction to finding, editing, and using data in Wikidata, focusing on people associated with paleontology collections (e.g., collectors, researchers, collections staff) as subjects. Together, workshop participants created or enhanced Wikidata records for around 100 individuals from a shared list, including a dozen female collectors previously known only by their husbands' names, e.g., "Mrs. Paul E. Drez," who we now know is Nancy Sue Drez. At the conclusion of the workshop, participants collaborated on an open-access document to share the findings. The goal of *Guidelines for Using Wikidata to Mobilize Information about People in Collections: A Paleontology Perspective* (see Bauer et al. 2022 for repository) is to further encourage uptake in the paleontological and collections communities.

The first workshop was focused almost exclusively on the people connected to paleontological collections. A follow-up workshop in October 2022 entitled, *Using Wikidata to capture and share information about paleontological collecting sites*, significantly broadens this scope by exploring how Wikidata can be used to mobilize community knowledge about specimen collecting events, such as sites or expeditions. Paleontological collecting sites pose an information management challenge because they tend to be detail rich, attached to local but stable identifier schemes, and associated with specimens curated in multiple institutions. Participants in this workshop established a defined scope of information associated with a paleontological collection site–including flagging aspects of the data that may be problematic to share publicly–and subsequently were able to assess how these information needs might, or might not, fit into a linked data model on Wikidata. These findings are particularly relevant to the global collections community where we see a trend towards sharing specimen data from an event-based perspective (Schindel and Cook 2018, Robertson et al. 2019), as has long been the norm for paleontological specimens.

Keywords

biodiversity knowledge graph, paleoinformatics, identifiers, people, specimen collecting event, linked data, semantic web

Presenting author

Holly Little

Presented at

TDWG 2022

References

- Bauer J, Burkhalter R, Karim T, Krimmel E, Landis M, Leachman S, Little H, Lorente M, Mills S, Neu-Yagle N, Norton B, Paul D, Shorthouse D, Utrup J, Van Veldhuizen J, Walker L (2022) Guidelines for Using Wikidata to Mobilize Information about People in Collections: A Paleontology Perspective. Zenodo. https://doi.org/10.5281/zenodo.6977243
- Groom Q, Güntsch A, Huybrechts P, Kearney N, Leachman S, Nicolson N, Page RDM, Shorthouse DP, Thessen AE, Haston E (2020) People are essential to linking biodiversity data. Database 2020 https://doi.org/10.1093/database/baaa072
- Güntsch A, Groom Q, Ernst M, Holetschek J, Plank A, Röpert D, Fichtmüller D, Shorthouse DP, Hyam R, Dillen M, Trekels M, Haston E, Rainer H (2021) A botanical demonstration of the potential of linking data using unique identifiers for people. PLOS ONE 16 (12). https://doi.org/10.1371/journal.pone.0261130
- Robertson T, Gonzalez M, Høfft M, Grosjean M (2019) Documenting Natural History Collections in GBIF. Biodiversity Information Science and Standards 3 (e37216). https://doi.org/10.3897/biss.3.37216
- Schindel D, Cook J (2018) The next generation of natural history collections. PLOS Biology 16 (7). https://doi.org/10.1371/journal.pbio.2006125